

月降水量に適した情報の与え方と 2変数最大エントロピー分布

寒川 典昭*・荒木 正夫**・清水 克彦***

Methods for Obtaining Suitable Information for Monthly Precipitation and 2-Variable Maximum Entropy Distributions

Noriaki SOGAWA*, Masao ARAKI**, and Katsuhiko SHIMIZU***

Synopsis

Methods for obtaining optimum information for a 2-variable maximum entropy distribution (2MED) were examined to estimate the hydrologic variable for the design of a water-supply system between two basins. Information consisting of statistical moments, and combined information of statistical moments and the expectations of an exponential function were tested with a set of rainfall data in Nagano, Nagoya and Tokyo. The goodness-of-fit (Log-likelihood, Akaike's information criterion and histogram) and stability of the return period were used as evaluation criteria in selecting suitable information. Though the $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$ distribution for combined information ranked first in terms of goodness-of-fit, it did not provide good stability of the return period because it exhibits a sensitive reaction with regards to removed data. The $2M(x, x^2, y, y^2, xy)$ distribution for statistical moments only showed excellent results for both evaluation criteria. Therefore, it was judged from these results that the $2M(x, x^2, y, y^2, xy)$ distribution was the optimum distribution for the data set used in this investigation.

1. ま え が き

渇水問題に対処する水工計画学的手法の1つに流域間導水計画がある。これは一方の流域に水量的に余裕がある時、他の地域に流域を越えて送水することにより、他の地域の水不足を解消しようとするものである。

さて、この計画が水量面からみて有効に機能し得るかどうかを検討するためには2地点間の水水量を確率変数とみなした2変数確率分布が必要となる。従来、水文統計における2変数分布としては、正規分布¹⁾、指数分布^{2),3),4)}、ガンマ分布^{5),6)}が用いられてきた。ところが、2変数正規分布は分布形が左右対称であり周辺分布が正規分布をする場合に、2変

* 社会開発工学科 助教授

** 熊本工業大学土木工学科 教授

*** 玉野総合コンサルタント(株)

数指数分布は周辺分布が指数分布をする場合に利用されるが、このような水量は特殊な場合に限られる。一方、2変数ガンマ分布は広く非対称形を表現することができ、応用面において有効な分布とされている。ところが、密度関数式に変形ベッセル関数を含んでいるためその取り扱いに困難を生じる場合がある²⁾とともに、今のところ一般形の最尤解が求められていない⁷⁾。また、3変数以上に拡張したときパラメータ同定が困難であるという欠点も有している。

そこで、我々は多くの水量に柔軟に対応できる分布として2変数最大エントロピー分布を提案し⁸⁾、統計モーメントを情報として母集団の再現性、及び実データへの適用性について詳細な検討を行った⁹⁾。この分布は、1変数最大エントロピー分布の性質を保持したままで2変数に拡張したものであるため、小標本を取り扱う水文統計にとって有効な分布¹⁰⁾であること、3変数以上の多変数分布¹¹⁾及び条件付き分布¹²⁾への拡張が容易に行えること、及び密度関数式が簡潔で利用しやすいことなどの特徴を持つことは、1変数の場合と同様である。

ところが、上述の2変数最大エントロピー分布は理論的には任意関数の期待値で情報を与える形で導出されているものの、シミュレーション及び実データへの適用の際は統計モーメントのみの情報に限定されている。しかし、最近の研究では、1、2変数の最大値水量を対象とするときには、統計モーメントと指数関数の期待値を組み合わせて情報とした場合の方が優れていることが明らかにされてきた^{13),14)}。本稿はこのことが流域間導水計画についても同様であるかどうかを、この計画を立案するときの基本的な水量の1つである月降水量を取り上げ、宝ら¹⁵⁾の開発した最適な分布の選定手順に従って検討したものである。

その内容を簡単に述べれば以下のようなになる。

第2章では、任意関数の期待値で情報を与えた場合の2変数最大エントロピー分布の確率密度関数を推定し、そのパラメータの同定を逐次計算法を用いて展開する。

第3章では、推定された分布のデータに対する全体的な適合性とリターンピリオドの安定性を評価するための規準について概説する。

第4章では、1例として長野、名古屋、東京の8月降水量を取り上げ、第3章で記述した規準を用いて2変数最大エントロピー分布への最適な情報の与え方を探索する。

最後に第5章では、本研究で得られた成果を要約してあとがきとするとともに、今後の検討課題についても触れておく。

2. 理 論 式

2.1 密度関数の推定

連続な2つの確率変数 X, Y に対する同時確率密度関数を $p(x, y)$ とすると、結合エントロピー $H(X, Y)$ は次式で与えられる¹⁶⁾。

$$H(X, Y) = -\iint_{R^2} p(x, y) \ln p(x, y) dx dy \quad (1)$$

なお、以下では「 \iint_{R^2} 」を「 \iint 」と表わす。

一方、確率密度関数が具備すべき条件は

$$\iint p(x, y) dx dy = 1 \quad (2)$$

である。また、適当な積分可能性をもつ任意の2変数関数 $g(s, t)$ と確率変数 X, Y に対して確率変数 $g(X, Y)$ の期待値を

$$\iint g(x, y) p(x, y) dx dy = E[g(x, y)] \quad (3)$$

で定義する。ただし、左辺の積分は存在するものと仮定する。ここに、 $E[\cdot]$ は期待値記号である。

いま、 $g(X, Y)$ として $g_r(X, Y) (r=1, 2, \dots, N)$ を採用し

$$E[g_r(X, Y)] = a_r, \quad r=1, 2, \dots, N \quad (4)$$

すなわち

$$\iint g_r(x, y) p(x, y) dx dy = a_r, \quad r=1, 2, \dots, N \quad (5)$$

が成立するものとする。

そこで、(2)、(5)式を制約条件として(1)式を最大にする $p(\cdot, \cdot)$ を求める。このような立場から求めた $p(\cdot, \cdot)$ を水文頻度分析に用いることの妥当性は、著者ら¹⁰⁾がすでに議論してきたとおりである。従って、この問題を $p(\cdot, \cdot)$ についてラグランジュの未定乗数法で解くと

$$p(x, y) = \exp\left\{-\lambda_0 - \sum_{r=1}^N \lambda_r g_r(x, y)\right\} \quad (6)$$

が得られる。ここに、 $\lambda_r (r=0, 1, 2, \dots, N)$ はラグランジュ乗数であり、推定分布 $p(\cdot, \cdot)$ からみれば、その分布のパラメータである。(6)式の右辺を $2M(g_1(\cdot, \cdot), g_2(\cdot, \cdot), \dots, g_N(\cdot, \cdot))$ と略記する。(6)式は一般的な2変数最大エントロピー分布(2-Variable Maximum Entropy Distribution, (2MED))であり、 $g_r(\cdot, \cdot)$ に上述した条件を満足する適当な関数を与えることにより、具体的な分布形が決定される。

2.2 パラメータ同定法

まず、(6)式を(2)式に代入し、 λ_0 について解くと次式となる。

$$\lambda_0 = \ln\left[\iint \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x, y)\right\} dx dy\right]. \quad (7)$$

次に、(6)式を(5)式に代入し、(7)式を用いて整理すると

$$\iint g_s(x, y) \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x, y)\right\} dx dy = a_s \iint \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x, y)\right\} dx dy, \quad s=1, 2, \dots, N \quad (8)$$

が得られる。従って、(6)式のパラメータ $\lambda_r (r=1, 2, \dots, N)$ は非線形方程式(8)を解くことにより決定される。 λ_0 は決定された $\lambda_r (r=1, 2, \dots, N)$ を(7)式に代入することにより算定される。こうして求まった λ_0 は $\lambda_r (r=1, 2, \dots, N)$ に依存しているため、自由パラメータとはみなされない。

ここで、(8)式を λ_r について解くために、Wragg・Dowson¹⁷⁾の方法を2変数に拡張して用いる。

まず、 λ_r の近似値として a_r を与え、残差を ε_r とおくと

$$\lambda_r = a_r + \varepsilon_r, \quad r = 1, 2, \dots, N \quad (9)$$

が成り立つ。そこで、上式を(8)式に代入して、 $\exp\{-\varepsilon_r g_r(x, y)\}$ をテーラー展開し、 $(\varepsilon_r)^2$ 以上の項を無視して整理すると

$$\begin{aligned} & \sum_{t=1}^N [\varepsilon_t \iint g_s(x, y) g_t(x, y) \exp\{-\sum_{r=1}^N a_r g_r(x, y)\} dx dy \\ & - a_s \iint g_t(x, y) \exp\{-\sum_{r=1}^N a_r g_r(x, y)\} dx dy] \\ & = \iint g_s(x, y) \exp\{-\sum_{r=1}^N a_r g_r(x, y)\} dx dy \\ & - a_s \iint \exp\{-\sum_{r=1}^N a_r g_r(x, y)\} dx dy, \quad s = 1, 2, \dots, N \end{aligned} \quad (10)$$

となる。ここで、 $g_0(x, y) = 1$, $a_0 = 1$ とし

$$c_{s,t} = \iint g_s(x, y) g_t(x, y) \exp\{-\sum_{r=1}^N a_r g_r(x, y)\} dx dy, \quad s, t = 0, 1, 2, \dots, N \quad (11)$$

とおくと、(10)式は

$$\sum_{t=1}^N \varepsilon_t \{c_{s,t} - a_s c_{0,t}\} = c_{s,0} - a_s c_{0,0}, \quad s = 1, 2, \dots, N \quad (12)$$

と書き改められる。さらに

$$\left. \begin{aligned} p_{s,t} &= c_{s,t} - a_s c_{0,t} \\ q_s &= c_{s,0} - a_s c_{0,0}, \quad s, t = 1, 2, \dots, N \end{aligned} \right\} \quad (13)$$

とおくと、(12)式は

$$\sum_{t=1}^N p_{s,t} \varepsilon_t = q_s, \quad s = 1, 2, \dots, N \quad (14)$$

となり、 ε_r に関する N 元連立1次方程式が得られる。

数値計算では、初期値 $\{a_1, a_2, \dots, a_N\}$ を与え、残差 $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N\}$ について(14)式を解く計算を残差が十分小さくなるまで繰り返すことにより、解 $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ に実用的にはほぼ等しい近似値を得ることができる。

3. 推定分布の評価規準

水文量の頻度分析では母集団の存在を仮定し、この分布形を未知の母集団から生じたデータから推定する。推定された分布が複数ある場合、最良の分布を選択しなければならない。宝ら¹⁹⁾は1変数の場合において、全体的な適合度が良く、しかも右裾が安定している分布が良いという立場から、推定分布の1つの評価手順を提案した。本稿でも基本的にはこれと同じ立場をとる。ただし、裾を評価するためのJackknife及びBootstrap推定

値と推定誤差の理論式¹⁸⁾は2変数の場合については導出されていないため、リターンピリオドの推定値とその安定性で代用する。以下に、ここで用いた分布形の評価規準を概説する。

3.1 全体的適合性の評価規準

(1) 対数尤度(LL)

2変数の場合のLLは

$$LL = \sum_{i=1}^n \ln f(x_i, y_i; \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k) \quad (15)$$

で表わされる。ここに、 $f(\cdot, \cdot)$ は確率密度関数、 x_i, y_i はデータ、 N はデータ数、 $\hat{\theta}_j (j = 1, 2, \dots, N)$ は推定パラメータである。

最尤法の観点からは、LLが大きい程優れた分布と判断される。

(2) 赤池の情報量規準(AIC)

パラメータ数の多い分布は、データに対する過剰適合、及び確率水文量の不安定性の問題を生じさせるとともに、パラメータ同定の計算時間が大きくなり、必ずしも最適な分布とは言えない。そのような問題を解決するために、赤池¹⁹⁾はパラメータ数も評価に取り込める規準としてAICを提案した。一般にAICは

$$AIC = -2 \times (\text{推定分布の最大対数尤度}) \\ + 2 \times (\text{推定分布の自由パラメータ数}) \quad (16)$$

と表現される。この値が小さなもの程良い分布である。

3.2 分布形の裾の部分の評価規準

ここでは分布形の裾の部分から評価されるリターンピリオドあるいは確率水文量の安定性規準について述べる。1変数の場合、宝ら¹⁵⁾はリサンプリング手法であるJackknife法とBootstrap法を用いているが、2変数の場合には今のところこれらの理論式が導出されていない。従って、ここでは次のようにしてリターンピリオドの安定性を評価する。

- 1) n 個のデータ(通常すべてのデータ)から i 番目のデータを除外したデータセットを n 組つくる。
- 2) n 組のデータセットを用いて、一定の情報の下でパラメータを同定し、異なったパラメータの値をもつ n 個の分布形を求める。
- 3) n 個の分布形から、ある設定値に対する n 個のリターンピリオドをそれぞれ算定し、その変動係数を計算する。
- 4) 情報の与え方を変えて、2), 3)を繰り返す。
- 5) 求めた変動係数の小さな値を示した分布が、裾の部分の安定した良い推定分布と判断する。

上述の手法は、Jackknife法、Bootstrap法等と同様に、リサンプリング手法であるが、リターンピリオドを評価対象にしている点からみれば、寒川ら¹³⁾の用いた手法に類似している。

3.3 最適な分布の選定

ここでは、宝ら¹⁵⁾が1変数確率分布の評価に用いた手順、及びその考え方を踏襲して最適な1変数最大エントロピー分布の選定を行った寒川ら¹³⁾の手順を参考にして、次のよう

に最適な分布を選定する。

まず、3.1で記述した LL 、 AIC により分布形の全体的な適合性を判定し、ある程度スクリーニングをする。次に、宝ら¹⁵⁾及び寒川ら¹³⁾の方法で用いられていないが、 LL 、 AIC で上位のものに限り確率分布のメッシュ確率を求め、それを3次元にグラフ化した体積柱状図を用いて全体的な適合性の判定を補う⁸⁾。最後に、全体的適合性の評価から選ばれた分布を用いて、3.2で述べた手順に従ってリターンピリオドの変動係数を計算し、その値が最も小さい分布を最適な分布と判断する。

なお、 $2MED$ の適合性評価に用いられる AIC は、厳密には AIC を計算したことになっていない。それは、 $2MED$ のパラメータ同定は分布形推定の一連の過程の中に含まれた最大エントロピー法であり、最尤法を用いていないためである。しかし、1変数でも行なわれている様に、他のモデルと比較する場合にこの値は有効となるので、ここでは近似的な意味で AIC を利用している¹³⁾。

4. 月降水量データへの適用結果

4.1 分布形の全体的な適合性

本章では、図1に示す長野、名古屋、東京の明治24年～昭和57年（92個）の8月降水量を採用した。地点間の組み合わせは、(a)長野(x)―名古屋(y)、(b)長野(x)―東京(y)である。な

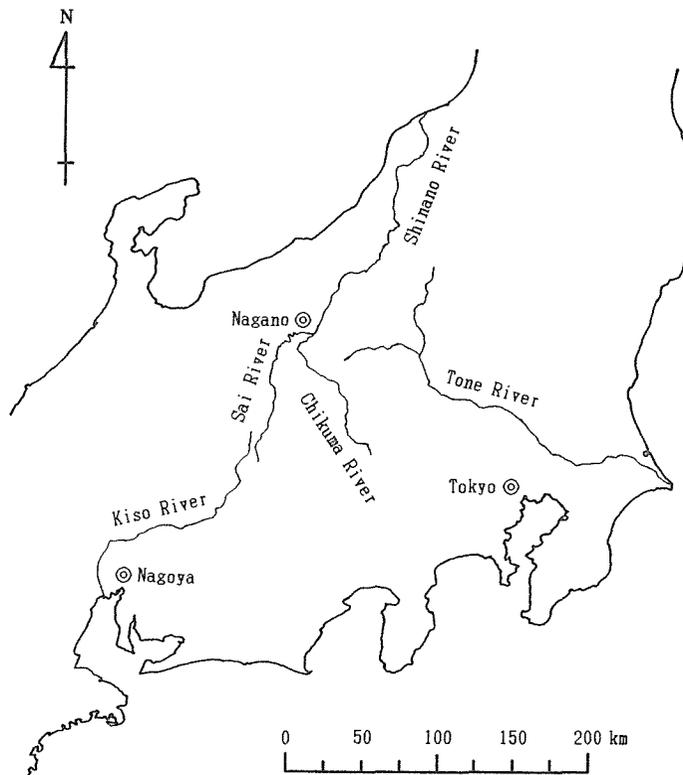


Fig. 1 Rain gauge location.

Table 1 Values of the LL and the AIC in the best five of the $2M(x^a, \exp(-bx/M_x), y^c, \exp(-dy/M_y), xy)$ and in the $2M(x, x^2, y, y^2, xy)$.

(a) Nagano-Nagoya

Ranking	$p(x, y)$				LL	AIC
	$2M(x^a, \exp(-bx/M_x), y^c, \exp(-dy/M_y), xy)$					
	a	b	c	d		
1	2	4	2	4	-1030.113	2070.226
2	2	3	2	4	-1030.232	2070.464
3	2	4	2	3	-1030.234	2070.468
4	2	3	2	3	-1030.353	2070.707
5	2	4	2	2	-1030.393	2070.785
39	$2M(x, x^2, y, y^2, xy)$				-1032.492	2974.983

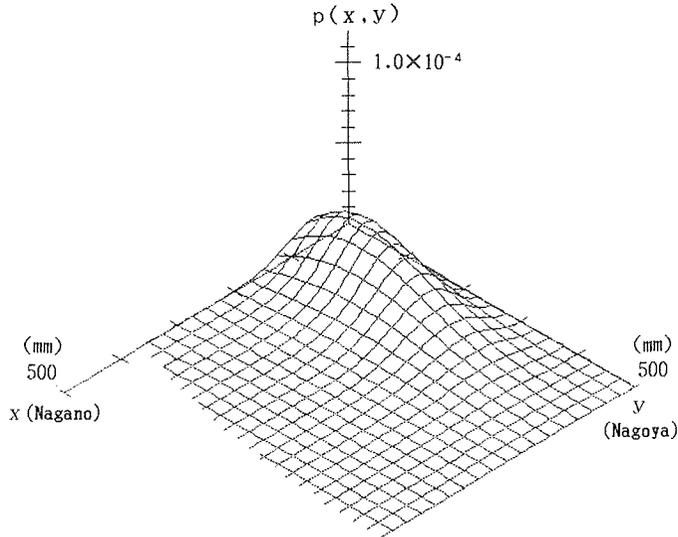
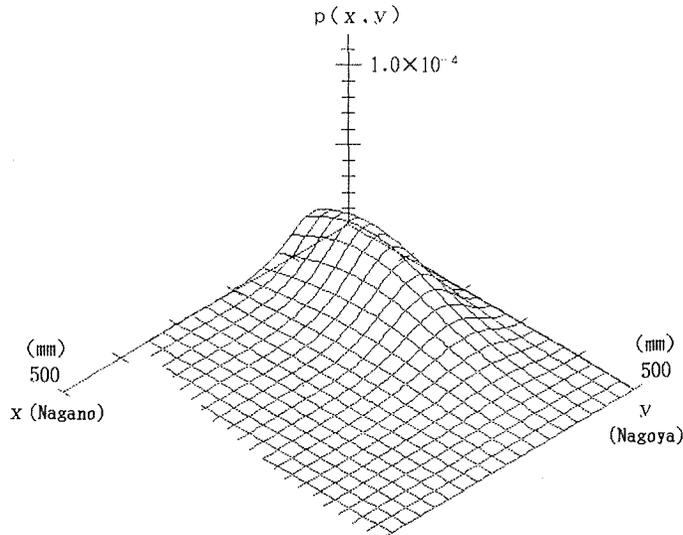
(b) Nagano-Tokyo

Ranking	$p(x, y)$				LL	AIC
	$2M(x^a, \exp(-bx/M_x), y^c, \exp(-dy/M_y), xy)$					
	a	b	c	d		
1	2	4	2	4	-1027.005	2064.010
2	2	3	2	4	-1027.078	2064.155
3	2	4	2	3	-1027.228	2064.456
4	2	3	2	3	-1027.302	2064.604
5	2	4	2	2	-1027.380	2064.761
78	$2M(x, x^2, y, y^2, xy)$				-1029.372	2068.744

(c) Nagoya-Tokyo

Ranking	$p(x, y)$				LL	AIC
	$2M(x^a, \exp(-bx/M_x), y^c, \exp(-dy/M_y), xy)$					
	a	b	c	d		
1	2	4	2	4	-1083.755	2177.510
2	2	3	2	4	-1083.874	2177.748
3	2	4	2	3	-1083.920	2177.839
4	2	4	2	2	-1084.018	2178.036
5	2	3	2	3	-1084.039	2178.077
12	$2M(x, x^2, y, y^2, xy)$				-1084.650	2179.300

お、ここで用いる2MEDは、治水計画問題への適用結果から類推して、パラメータ数を5個に固定した $2M(x, x^2, y, y^2, xy)$ と $2M(x^a, \exp(-bx/M_x), y^c, \exp(-dy/M_y), xy)$ の2つのタイプである¹⁴⁾。ただし、 M_x, M_y はそれぞれ確率変数 x, y に対応するデータの平均値であり、 a, b, c, d は1~4の整数とする¹⁴⁾。 $2M(x, x^2, y, y^2, xy)$ は2次まで

(a) $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$ (b) $2M(x^2, \exp(-4x/M_x), y^2, \exp(-2y/M_y), xy)$ Fig. 2 Curved surface of $2M(x^a, \exp(-bx/M_x), y^c, \exp(-dy/M_y), xy)$ in Nagano-Nagoya.

の統計モーメントのみを情報とした分布であり、 $2M(x^a, \exp(-bx/M_x), y^c, \exp(-dy/M_y), xy)$ は統計モーメントと指数関数の期待値を組み合わせて情報とした分布である。

(a) 長野一名古屋

組み合わせた情報をもつ $2MED$ は、 $a \sim d$ のすべての組み合わせについて計算した結果、179パターンでパラメータが同定できた。表1(a)のRanking1~5にこれらの分布の上位5パターンの LL と AIC の値を示している。パラメータ数を5個に固定しているため、 LL と AIC は同じ順位を与えている。これより、 a, c は2で、 b, d は4で情報を与

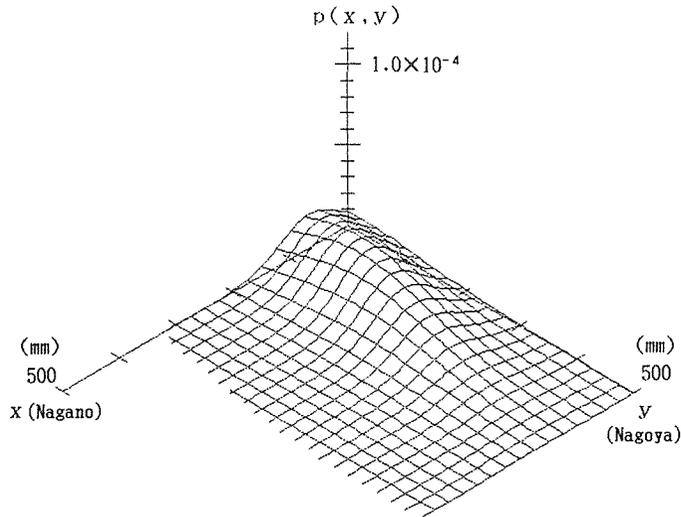


Fig. 3 Curved surface of $2M(x, x^2, y, y^2, xy)$ in Nagano-Nagoya.

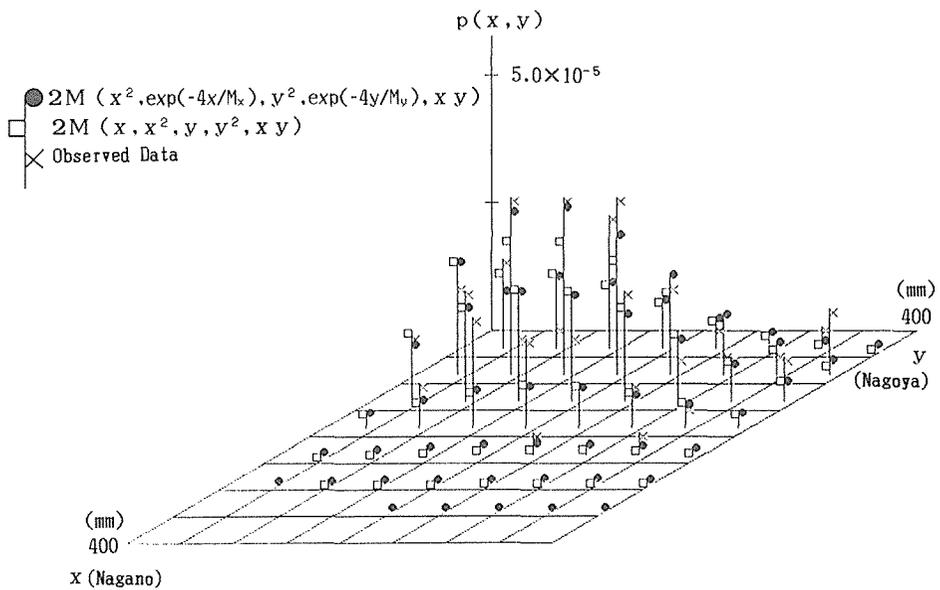
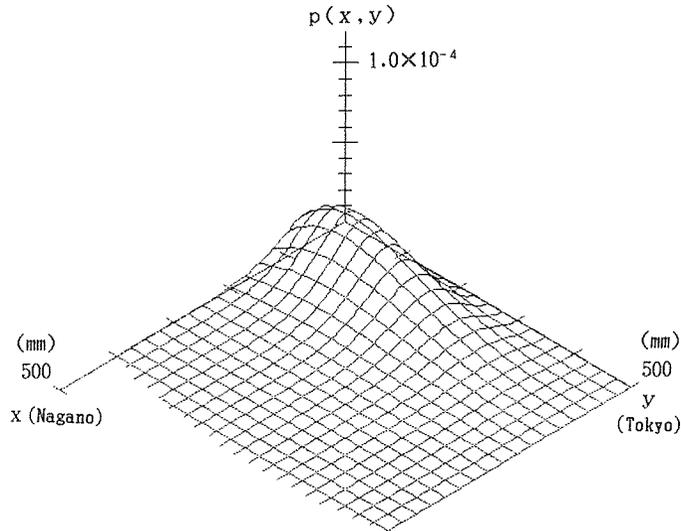
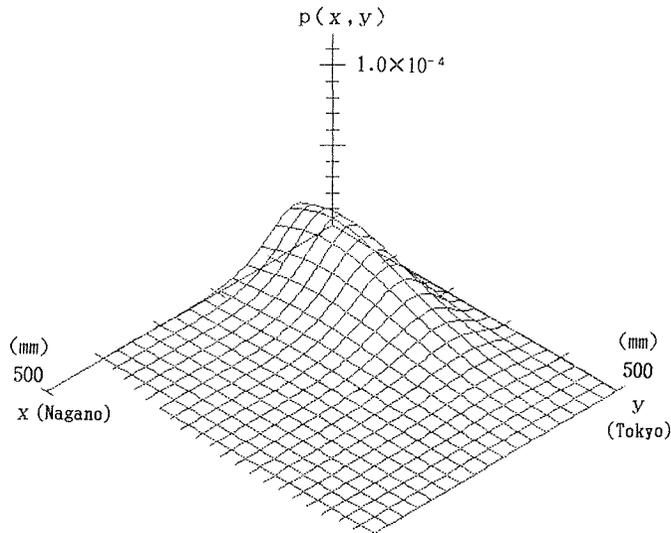


Fig. 4 Histogram of $2M(x^2, \exp(-4x / M_x), y^2, \exp(-4y / M_y), xy)$, $2M(x, x^2, y, y^2, xy)$ and observed data in Nagano-Nagoya.

えた $2M(x^2, \exp(-4x / M_x), y^2, \exp(-4y / M_y), xy)$ が最も良い適合性を示すことがわかる。また、上位の分布は、a, cは2で、b, dは高次で情報を与えており、このような情報の与え方が全体的適合性に良い評価を与えることになる。図2には、1位と5位の密度関数立体図を掲載している。

表1(a)のRanking39と図3は統計モーメントのみで情報を与えた $2M(x, x^2, y, y^2, xy)$

(a) $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$ (b) $2M(x^2, \exp(-4x/M_x), y^2, \exp(-2y/M_y), xy)$ Fig. 5 Curved surface of $2M(x^a, \exp(-bx/M_x), y^c, \exp(-dy/M_y), xy)$ in Nagano-Tokyo.

の LL , AIC の値, 及び密度関数立体図である. この LL , AIC は上述の179パターンと比較すると39位に相当するが, 上位の分布との間に大差はない. また, 図3の密度関数立体図の形状は, 全体的にみれば図2に類似している.

図4は $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$ と $2M(x, x^2, y, y^2, xy)$ の実データに対する適合性を視覚的にみた体積柱状図である. 64個のメッシュのうち, 前者の分布の方が実データに近似しているメッシュ数は3個, その逆も3個であり, 残りのメッシュでは2つの分布の実データへの近似度の良否は判断し難かった.

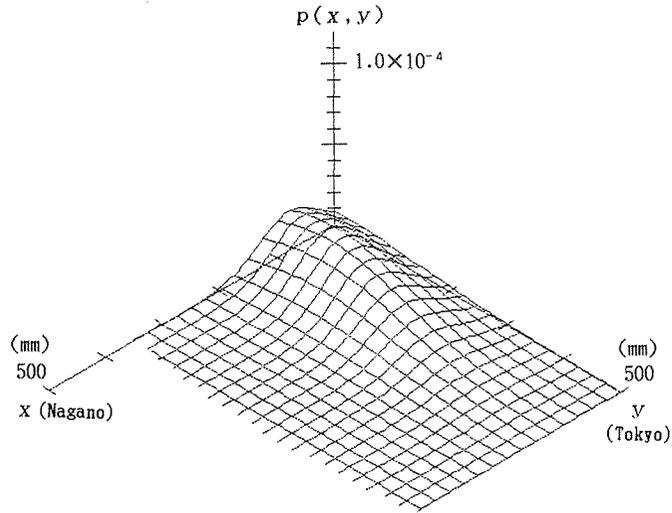


Fig. 6 Curved surface of $2M(x, x^2, y, y^2, xy)$ in Nagano-Tokyo.

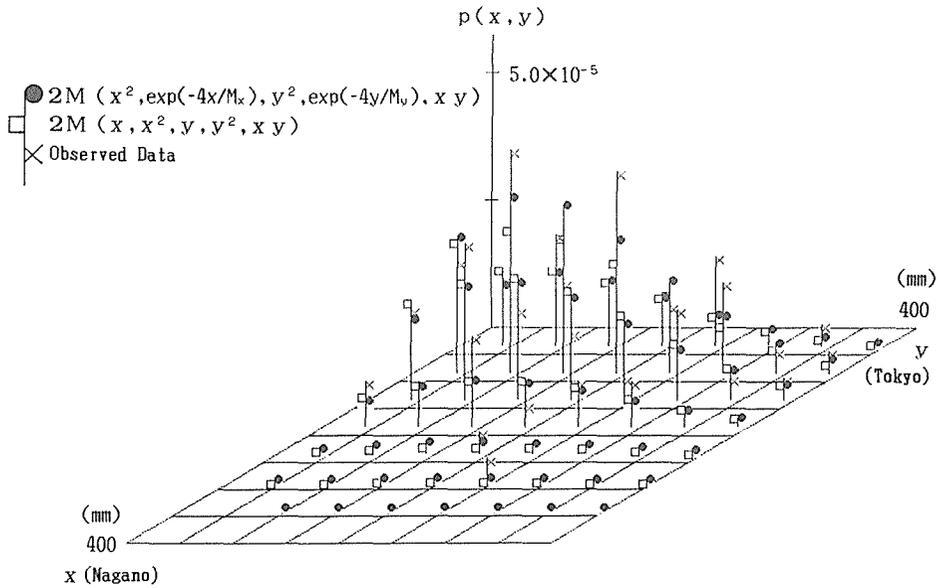


Fig. 7 Histogram of $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$, $2M(x, x^2, y, y^2, xy)$ and observed data in Nagano-Tokyo.

(b) 長野—東京

ここでも、組み合わせた情報をもつ $2MED$ について、 $a \sim d$ のすべての組み合わせを計算した結果、177パターンでパラメータが同定できた。表 1 (b) の Ranking 1~5 にこれらの分布の上位 5 パターンの LL と AIC の値を示している。ここでも a, c は 2 で、 b, d は高次で情報を与えた分布が良い結果を与えており、 $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$ が最良の分布となった。また、1~5 位の情報の与え方は長野一名

古屋の場合と全く同じである。図5には、1位と5位の密度関数立体図を掲載している。

表1(b)のRanking78と図6は統計モーメントで情報を与えた $2M(x, x^2, y, y^2, xy)$ の LL, AIC の値、及び密度関数立体図である。この LL, AIC は上述の177パターンと比較すると78位に相当するが、ここでも上位の分布との間に大差はない。また、長野一名古屋の場合と同様に、図6の密度関数立体図の形状は、全体的には図5に類似している。

図7は図4と同様に、 $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy), 2M(x, x^2, y, y^2, xy)$ 、及び実データの体積柱状図を描いたものである。64個のメッシュのうち、前者の分布の方が実データに近似しているメッシュ数は2個であり、残りのメッシュでは2つの分布の実データへの近似度の良否は判断し難かった。

なお、ここでは紙面の都合で図は割愛するが、参考までに表1(c)に名古屋一東京で計算した LL と AIC の値も掲載している。上位3位の分布は長野一名古屋、長野一東京の場合に同じであり、統計モーメントのみで情報を与えた分布は組み合わせた情報をもつ155パターンの分布中12位であったが、それより上位の分布との間で全体的な適合性において大差はなかった。

4.2 リターンピリオドの安定性

4.1では、どの地点間の組み合わせでも上位3位の分布は同じであったため、ここでは長野一東京の場合に限ってリターンピリオドの安定性を検討した。また、本稿での問題は、渇水に対処するための利水計画に適した分布を探索することにあるため、この問題で対象となる x, y ともに一定の設定値を超過しないリターンピリオドの安定性、すなわち x, y ともに原点に近い裾の部分の安定性をみることにした。今回の設定値は、4.1(b)の LL, AIC 1位の分布でリターンピリオド100年付近となる x, y の平均値の(1/3.198)の値($x:31.6345$ mm, $y:47.0495$ mm)と、同様に50年付近となる x, y の平均値の(1/2.613)の値($x:38.7164$ mm, $y:47.0495$ mm)である。前者を(a)、後者を(b)とする。比較に用いた分布は、 LL, AIC で最も良い値を示した $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$ と、1変数の場合から原点の方に尾を引くことが類推される統計モーメントのみで情報を与えた $2M(x, x^2, y, y^2, xy)$ の2つの分布である。表2には設定した値に対するリターンピ

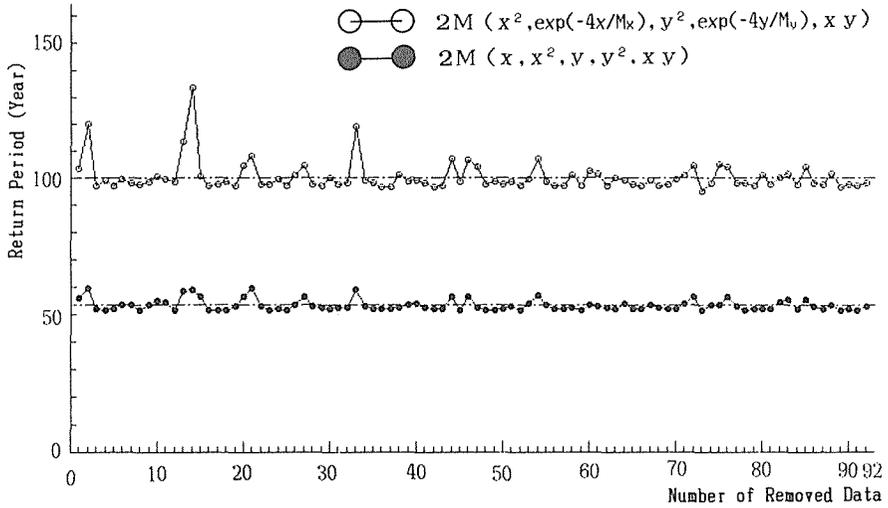
Table 2 Coefficient of variation of the return period.

(a) Case: $x=31.6345$ (mm), $y=47.0495$ (mm)

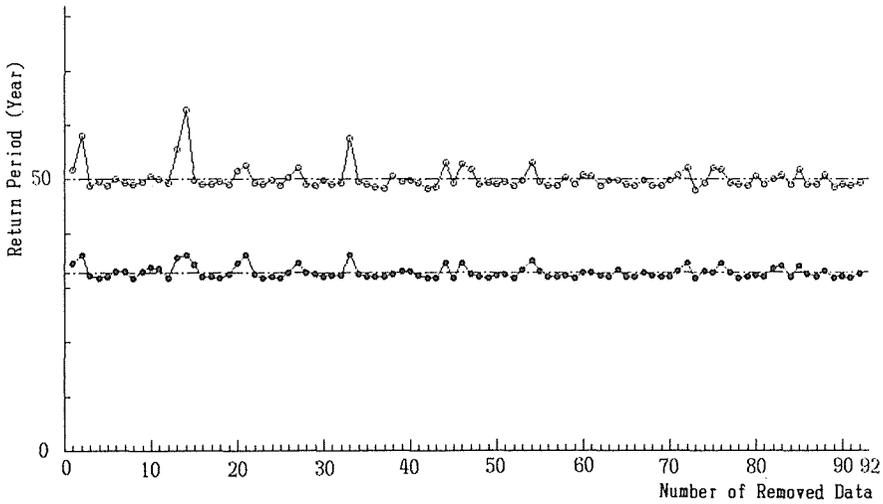
$p(x, y)$	Mean	Unbiased Variance	Sample Variance	Coefficient of Variation
$2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$	100.246	30.6077	30.2750	0.0548879
$2M(x, x^2, y, y^2, xy)$	53.323	4.05075	4.00672	0.0375388

(b) Case: $x=38.7164$ (mm), $y=57.5824$ (mm)

$p(x, y)$	Mean	Unbiased Variance	Sample Variance	Coefficient of Variation
$2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$	50.0848	4.84009	4.78748	0.436865
$2M(x, x^2, y, y^2, xy)$	32.7700	1.25663	1.24297	0.340216



(a) Case: $x = 31.6345(\text{mm})$, $y = 47.0495(\text{mm})$



(b) Case: $x = 38.7164(\text{mm})$, $y = 57.5824(\text{mm})$

Fig. 8 Stability of the return period.

リオドの平均値、不偏分散、標本分散、及び変動係数を示している。これより、(a)では、リターンピリオドの平均値と変動係数が、組み合わせた情報をもつ方で100年、0.055に対し、統計モーメントのみの情報では53年、0.038であるため、後者の方が原点方向に尾を引き、安定していることがわかる。また、(b)の場合も、前者が50年、0.437に対し、後者は33年、0.340となり(a)と同様の傾向を示している。図8は抜いたデータ番号とリターンピリオドとの関係、すなわちリターンピリオドの変動を示したものである。抜いたデータの値が小さな場合は大きなリターンピリオドを与えており、その傾向は組み合わせた情報をもつ場合に著しい。以上より、リターンピリオドの安定性の検討では、 $2M(x, x^2, y, y^2, xy)$ が最適と判断される。

従って、4.1, 4.2より総合的に判断すると、 $2M(x, x^2, y, y^2, xy)$ は、全体的な適合性では LL, AIC の上位の分布に大きく劣ることなく、利水計画上重要な確率変数の小さい値に対する非超過確率から計算されるリターンピリオドが安定していることより、この分布を最適とみなすことにする。

5. あとがき

本稿では、長野、名古屋、東京の8月降水量に対して、客観的な評価規準を用いて水量の母集団を実務レベルで満足される精度で推定し、 $2MED$ への流域間導水計画問題に適した情報の与え方を探索した。実データへの適用から得られた成果を要約すると次のようになる。

- 1) LL, AIC からデータに対する分布形の全体的な適合性を検討した。その結果、長野—名古屋、長野—東京、名古屋—東京のいずれの場合においても、統計モーメントによる情報は、 x, y ともに2次で、指数関数による情報は高次で組み合わせた分布が良いと判断できた。しかし、 $2M(x, x^2, y, y^2, xy)$ も組み合わせた情報をもつ分布と比較して大きく劣ることはなかった。
- 2) LL, AIC からみると、3つの地点間の組み合わせのどの場合でも上位にくる分布が類似しているため、長野—東京の場合に限って流域間導水計画上重要な原点に近い方向の非超過確率に対するリターンピリオドの安定性を評価した。ここでは、 $2M(x, x^2, y, y^2, xy)$ の分布が良く、 $2M(x^2, \exp(-4x/M_x), y^2, \exp(-4y/M_y), xy)$ との間には明確な差があった。
- 3) 1), 2)より、ここで用いたデータセットの場合、全体的に比較的良好な適合性を示し、しかも流域間導水計画上重要な部分のリターンピリオドが安定している $2M(x, x^2, y, y^2, xy)$ を優れた分布として選定した。

以上、3地点の月降水量を用いた計算結果について取りまとめたが、現段階では一般的な場合として最適な情報の与え方を論じるには計算例が少なく、得られた結果に偶然性を含んでいる可能性も考えられる。従って、今後は次の2点について検討したい。

- 1) 仮定した母集団からのシミュレーションデータを用い母集団に対する適合性及びリターンピリオドの安定性を評価することにより、母集団ごとに $2MED$ への最適な情報の与え方を検討する。
- 2) 多くの地点の実データに $2MED$ を適用し、地点の組み合わせを変えても上述の結果が得られるかどうかを検討する。

なお、本研究の数値計算では滝本信春君（住友建設㈱，当時信州大学工学部生）に御協力頂いた。記して謝意を表する。

参 考 文 献

- 1) 石原英雄，長尾正志：計画高水流量の二変数確率論的研究，京都大学防災研究所年報，第6号，pp. 181~188 (1963年)。
- 2) 端野道夫，杉 雄司：複数の二変数指数型分布の結合とその応用に関する研究，徳島大学工学部研究報告，第29号，pp. 49~57 (1984年)。

- 3) 端野道夫: Marked point process に伴う水文二変量の同時確率, 第28回水理講演会論文集, pp. 391~396 (1984年).
- 4) 端野道夫: Freund 分布による条件付確率降雨波形と計画降雨波形の決定法, 第31回水理講演会論文集, pp. 203~208 (1987年).
- 5) 長尾正志, 角屋 睦: 二変数ガンマ分布とその適用に関する研究(1)—二変数指数分布の基礎理論一, 京都大学防災研究所年報, 第13号, B, pp. 105~133 (1970年).
- 6) 長尾正志, 角屋 睦: 二変数ガンマ分布とその適用に関する研究(2)—二変数ガンマ分布の母数推定法一, 京都大学防災研究所年報, 第14号, B, pp. 43~56 (1971年).
- 7) 建設省近畿地方建設局: 河川工学百年の歩みと淀川, pp. 89~91, 社団法人近畿建設協会, 1978年.
- 8) 寒川典昭, 荒木正夫, 寺島 彰: 2変数 MEP 分布とその特性に関する研究, 第28回水理講演会論文集, pp. 397~402 (1984年).
- 9) 寒川典昭, 荒木正夫, 寺島 彰: 2変数最大エントロピー分布の適用性, 信州大学工学部紀要, 第62号, pp. 33~48 (1987年).
- 10) 寒川典昭, 荒木正夫: 水文事象の頻度分析への MEP 導入について, 土木学会論文報告集, 第335号, pp. 89~95 (1983年).
- 11) 寒川典昭, 荒木正夫, 佐藤健次: 多変数最大エントロピー分布とその基礎特性に関する研究, 土木学会論文集, 第375号/II-6, pp. 89~98 (1986年).
- 12) Sogawa, N., M. Araki, and T. Imai: Studies on multivariate conditional maximum entropy distribution and its characteristics, Journal of Hydrosience and Hydraulic Engineering, Vol. 4, No. 1, pp. 79~97 (1986).
- 13) 寒川典昭, 荒木正夫, 船橋太道: 年最大降雨に適した情報と母数の数と最大エントロピー分布, 土木計画学研究・論文集, 第7号, pp. 91~98 (1989年).
- 14) Sogawa, N., M. Araki and K. Shimizu: Information suitable for annual flood and 2-variate maximum entropy distribution, Pacific International Seminar on Water Resources Systems, pp. 103~116 (1989).
- 15) 宝 馨, 高棟琢馬: 水文頻度解析における確率分布モデルの評価規準, 土木学会論文集, 第393号/II-9, pp. 151~160 (1988年).
- 16) 例えば, 大泉充郎, 本多波雄, 野口正一: 情報理論, p. 78, オーム社, 1962年.
- 17) Wragg, A. and D. C. Dowson: Fitting continuous probability density functions over $[0, \infty)$ using information ideas, IEEE Transactions on Information Theory, Vol. IT-16, No. 2, pp. 226~230 (1970).
- 18) Efron, B.: The jackknife, the bootstrap and other resampling plans, SIAM Monograph, No. 38 (1982).
- 19) Akaike, H.: A new look at the statistical model identification, IEEE Transactions on Automatic Control, Vol. AC-19, No. 6, pp. 716~723 (1974).